

1 GGCCTGCCG CGCGGGGAGA AAGAAGCCGC GCCCCAGCCCG GCGTCCCAG
51 CAGCGCAGGG GAGGATCCCC CGCGAGTGCAC CCGGGAGCCA CCACAGACTC
101 TGGGAGGCTC GGCGGCTGGA GCAGCAGGCA GCTCCCCGCA GCTCCCGCG
151 CTTCCAGGCA GCTCTCTGAG CCGTGCCAGA GGCCCCGGCCC GCCATTCCA
201 GCCCCGAGCC ATGATGAAGA CTTTGTCCAG CGGGAACTGC ACGCTCAGTG
251 TGCCCCGCAA AAACATCATAAC CGCATGGTGC TGCTGGTGC CTCTCGGTG
301 GGCAAGAGCT CCATCGTGT CCGCTTCCTC AATGGCCGCT TTGAGGACCA
351 GTACACACCC ACCATCGAGG ACTTCCACCG TAAGGTATAC AACATCCGCG
401 GCGACATGTA CCAGCTCGAC ATCCTGGATA CCTCTGGCAA CCACCCCTTC
451 CCCGCCATGC GCAGGCTGTC CATCCTCACCA GGGGATGTC TCATCCTGGT
501 GTTCAGCCTG GATAACCGGG AGTCCTTCGA TGAGGTCAAG CGCCCTCAGA
551 AGCAGATCCT GGAGGTCAAG TCCTGCCTGA AGAACAAAGAC CAAGGGAGGCG
601 GCGGAGCTGC CCATGGTCAT CTGTGGCAAC AAGAACCGACC ACGGCAGAGCT
651 GTGCCGCCAG GTGCCCACCA CCGAGGCCGA GCTGCTGGTG TCGGGCGACG
701 AGAACTCCGC CTACTTCGAG GTGCGGCCA AGAAGAACAC CAACGTGGAC
751 GAGATGTCTC ACGTGCTCTT CAGCATGGCC AAGCTGCCAC ACGAGATGAG
801 CCCCCGCTTG CATCGCAAGA TCTCCGTGCA GTACGGTGCAC GCCTTCCACC
851 CCAGGCCCTT CTGCATGCGC CGCGTCAAGG AGATGGACGC CTATGGCATG
901 GTCTCGCCCT TCGCCGCCCG CCCCAGCGTC AACAGTGACC TCAAGTACAT
951 CAAGGCCAAG GTCCCTCGGG AAGGCCAGGG CCGTGAGAGG GACAAGTGCA
1001 CCATCCAGTG AGCGAGGGAT GCTGGGGCGG GGCTGGCCA GTGCCCTTCAG
1051 GGAGGTGCC CAGATGCC ACTGTGCGCA TCTCCCCACC GAGGCCCCGG
1101 CAGCAGCTTT GTTCACAGAC CTTAGGCACC AGACTGGAGG CCCCCGGCG
1151 CTGGCCTCCG CACATTCGTC TGCCTCTCA CAGCTTCCCT GAGTCCGCTT
1201 GTCCACAGCT CCTTGGTGGT TTCACTCTCTT CTGTGGGAGG ACACATCTCT
1251 GCAGCCTCAA GAGTTAGGCA GAGACTCAAG TTACACCTTC CTCTCCTGGG
1301 GTTGAAGAA ATGTTGATGC CAGAGGGGT AGGATTGTC CGTCATATGG
1351 AGCCTCTGG GACAAGCCTC AGGATGAAAAA GGACACAGAA GGCAGATGA
1401 GAAAGGTCTC CTCTCTCTG GCATAAACACC CAGCTTGGTT TGGGTGGCAG
1451 CTGGGAGAAC TTCTCTCCCA GCCCTGCAAC TCTTACGCTC TGTTCAAGCT
1501 GCCTCTGCAC CCCCTCCCAC CCCCAGCACA CACACAAGTT GCCCCCCCAGC
1551 TGCGCCTGAC ATTGAGCCAG TGGACTCTGT GTCTGAAGGG GGCCTGGCCA
1601 CACCTCTTAG ACCACGCCCA CCACTTAGAC CACGCCACC TCCTGACCAC
1651 GTTCTCTAGC CTCCTCTCTT AGGTCCCTCC GCCCCACAGT TGTGTTTGT
1701 TGTGGTTGCA GCTGTTTCG TGTCACTGAT AGTAGTAGAA ATGAAATCA
1751 TTGTACTGTA AAAGCCTAGT GACTCCCTCC TTGGCCAGGC CCTCACCCAG
1801 TTCAGATCCA CGGCCTCCAC CCGGGACGCC TTCCCTCTCT GCTCCCAAAC
1851 AGGGTTCCG TGGCCTAGT GCAGCTAGAC ATTGACCTCC GCCATTGAGC
1901 TCCACGGTT ACAGACAATT GCACAAGCGT GGGGTTGGCA GGCAGGACT
1951 GCTTTTTTTT AATGCTCCCA TTTCACAGAG GATACCCACCG AGACTCGGAG
2001 GGGACACGAT GAGCACCAAGG CCCCACCTTT GTCCCTAGC AAATTCAAGGG
2051 TACAGCTCCA CCTAGAACCA GGCTGCCCTC TACTGTGCTC GTTCCTCAAG
2101 CATTATTAA GCACCTACTG GGTGCTGGGT TCACTGTGTC CTAGGAAACC
2151 AAGAGGGTCC CCAGTCCTGG CCTCTGCCCG CCCCTGCTGC CCCACCACCT
2201 TCTGCACACA CAGCGGTGGG GAGGCAGGGGG GGAGCAGCTG GGACCCAGAA
2251 CTGAGCTGG GAGGGATCCG ACAGAAAAGC TCAGGGCGGG TCTTCTCCTT
2301 GTGCCGCCAG TTGGGCTATG CTGGGTACCA CCATGTACTC AGGCATGGTG
2351 GGTTTGAAC CCATAAACCA AAGGCCCTTG TCATCAGCTC TTAACAAGTA
2401 TATTTGTAT TTTAATCTCT CTAAACATAT TGAAGTTTA GGGCCCTAAG
2451 GAACCTTAGT GATCTTCTAT TGGGTCTTT TGAGGTTAG AGAGGGTAAG
2501 TAACTCCTC CAGGTCACAC AGCAAGCTG TGGGTGGCAG AAGCAAGCTA
2551 GCGCTGGCA TTCAGTACAT ACCACGATGT GCTCCCTCTC TTGATGCTTG
2601 GCCCCGGGG CCTTCAGGGC TTTGGGACAT CTTGCTCTCA ACCCTCTCCC
2651 TAGATCAGTC TGTGAGGGTC CCTGTAGATA TTGTGTACAC CATGCCCATG
2701 TATATACAAAG TACACACAGA TGTACACACA GATGTACACA TGCTCCAGCC
2751 CCAGCTCTGC ATACCTGCAC CTGCACCCCA GCCTTGGCCC CTGCTGGGT
2801 CTGTGCTCAA AGCAGCAGCT CCAACCCCTGC CTCTGTCCCC TTCCCCACCC
2851 ACTGCCTGAG CCTTCTGAGC AGACCAGGTA CCTTGGCTGC ACCGGTGTGT
2901 GGGCCGCTCT CACCCAGGCA CAGCCCCGCC ACCATGGATC TCCGTGTACA
2951 CTATCAATAA AAGTGGTTT GTTACAAAAA AAAAAAAA AAAAAAAA AAAAAAAA
3001 AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA AAAAAAAA
3051 AAAAAAAA AAAAAAAA AAAAAAAA AA

FIGURE 1, page 1 of 3

FEATURES:

5'UTR: 1-210
Start Codon: 211
Stop Codon: 1009
3'UTR: 1012

5' UTR ANALYSIS:

Query=cDNA clone
Sbjct: genomic sequence

5' UTR Exon 1, non-coding

Score = 58.0 bits (29), Expect = 2e-10
Identities = 29/29 (100%)
Strand = Plus / Plus

Query: 1 ggcgtcgccgcgcggggagaaagaagccg 29
|||||||||||||||||||||||||
Sbjct: 1535 ggcgtcgccgcgcggggagaaagaagccg 1563

5' UTR Exon 2, non-coding

Score = 349 bits (176), Expect = 3e-98
Identities = 176/176 (100%)
Strand = Plus / Plus

Query: 26 gccgcgcccagcccggtcccggcagcgcgcaggggaggatccccgcgcagtgaccggg 85
|||||||||||||||||||||||||
Sbjct: 2001 gccgcgcccagcccggtcccggcagcgcgcaggggaggatccccgcgcagtgaccggg 2060

Query: 86 agccaccacagactctgggaggctcgccggctggagcagcaggcagctccgcagctcc 145
|||||||||||||||||||||||||
Sbjct: 2061 agccaccacagactctgggaggctcgccggctggagcagcaggcagctccgcagctcc 2120

Query: 146 cggcggttccaggcagctctgtagccgtgccagaggccggccgcattccag 201
|||||||||||||||||||||||||
Sbjct: 2121 cggcggttccaggcagctctgtagccgtgccagaggccggccgcattccag 2176

Score = 563 bits (284), Expect = e-162
Identities = 284/284 (100%)
Strand = Plus / Plus

5' UTR Exon 3, Protein coding region begins at nucleotide 211

Query: 199 cagccccgagccatgtaaagactttgtccagcggaaactgcacgctcagtgtgcccgcc 258
|||||||||||||||||||||||||
Sbjct: 7474 cagccccgagccatgtaaagactttgtccagcggaaactgcacgctcagtgtgcccgcc 7533

Homologous proteins:**Top BLAST Hits**

	Score	E
CRA 18000005194969 /altid=gi 10047088 /def=ref NP_055125.1 sim...	538	e-152
CRA 18000005238449 /altid=gi 5059122 /def=gb AAD38928.1 AF13440...	512	e-144
CRA 332000009620725 /altid=gi 8118457 /def=gb AAF72997.1 AF2620...	342	2e-93
CRA 18000005232775 /altid=gi 7706359 /def=ref NP_057168.1 ras-...	342	2e-93
CRA 8700000006130 /altid=gi 7230768 /def=gb AAF43090.1 AF23915...	342	3e-93
CRA 18000005090459 /altid=gi 6677673 /def=ref NP_033052.1 RAS,...	341	6e-93
CRA 89000000197633 /altid=gi 7295299 /def=gb AAF50620.1 (AE003...	228	8e-59
CRA 105000014645240 /altid=gi 10503969 /def=gb AAG17979.1 AF177...	214	1e-54

BLAST dbEST hits:

gi 9345313 /dataset=dbest /taxon=960...	880	0.0
gi 9335874 /dataset=dbest /taxon=960...	846	0.0
gi 10143211 /dataset=dbest /taxon=96...	821	0.0
gi 9335309 /dataset=dbest /taxon=960...	813	0.0
gi 9150610 /dataset=dbest /taxon=9606...	662	0.0
gi 10144589 /dataset=dbest /taxon=96...	617	e-174
gi 9333908 /dataset=dbest /taxon=960...	599	e-169

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

Expression information from BLAST dbEST hits:

gi|9345313 Placenta choriocarcinoma
gi|9335874 Uterus endometrium adenocarcinoma
gi|10143211 Skin melanotic melanoma
gi|9335309 Uterus endometrium adenocarcinoma
gi|9150610 Skin melanotic melanoma
gi|10144589 Skin melanotic melanoma
gi|9333908 Uterus endometrium adenocarcinoma

Expression information from PCR-based tissue screening panels:

Human fetal whole brain

1 MMKTLSSGNC TLSVPAKNSY RMVVLGASRV GKSSIVSRFL NGRFEDQYTP
51 TIEDFHRKVV NIRGDMYQLD ILDTSGNHPF PAMRRLSILT GDVFILVFSL
101 DNRESFDEVK RLQKQILEVK SCLKNKTKEA AELPMVICGN KNDHGELCRQ
151 VPTTEAELLV SGDENSAYFE VSAKKNTNVD EMFYVLFSMA KLPHEMSPAL
201 HRKISVQYGD AFHPPRFICMR RVKEMDAYGM VSPFARRPSV NSDLKYIKAK
251 VLREGQARER DKCTIQ

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 2

1 9-12 NCTL
2 125-128 NTKK

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE

cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 4

1 84-87 RRLS
2 174-177 KKNT
3 202-205 RKIS
4 236-239 RRPS

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE

Protein kinase C phosphorylation site

Number of matches: 2

1 19-21 SYR
2 172-174 SAK

[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE

Casein kinase II phosphorylation site

Number of matches: 5

1 51-54 TIED
2 105-108 SFDE
3 154-157 TEAE
4 161-164 SGDE
5 177-180 TNVD

[5] PDOC00007 PS00007 TYR_PHOSPHO_SITE

Tyrosine kinase phosphorylation site

220-228 RRVKEMDAY

[6] PDOC00266 PS00294 PRENYLATION

Prenyl group binding site (CAAX box)

263-266 CTIQ

[7] PDOC00016 PS00016 RGD

Cell attachment sequence

63-65 RGD

[8] PDOC00017 PS00017 ATP_GTP_A

ATP/GTP-binding site motif A (P-loop)

26-33 GASRVGKS

BLAST Alignment to Top Hit:

>CRA|18000005194969 /altid=gi|10047088 /def=ref|NP_055125.1| similar
to mouse Ras, dexamethasone-induced 1; tumor endothelial
marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=278
Length = 278

Score = 538 bits (1372), Expect = e-152
Identities = 265/266 (99%), Positives = 265/266 (99%)

Query: 1 MMKTLSSGNCTLSPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY 60
MMKTLSSGNCTLSPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY
Sbjct: 13 MMKTLSSGNCTLSPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY 72
Query: 61 NIRGDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 120
NIRGDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK
Sbjct: 73 NIRGDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 132
Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVSAKKNTNVD 180
SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDEN AYFEVSAKKNTNVD
Sbjct: 133 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENCAYFEVSAKKNTNVD 192
Query: 181 EMFYVLFSMAKLPHEMSPALHRKISVQYDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 240
EMFYVLFSMAKLPHEMSPALHRKISVQYDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV
Sbjct: 193 EMFYVLFSMAKLPHEMSPALHRKISVQYDAFHPRPFCMRRVKEMDAYGMVSPFARRPSV 252
Query: 241 NSDLKYIKAKVLREGQARERDKCTIQ 266
NSDLKYIKAKVLREGQARERDKCTIQ
Sbjct: 253 NSDLKYIKAKVLREGQARERDKCTIQ 278

Complete Amino Acid Sequence of gi|10047088 /def=ref|NP_055125.1| similar
to mouse Ras, dexamethasone-induced 1; tumor endothelial
marker 2 [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=278
Length = 278

NOTE: UNDERLINED RESIDUES NOT PRESENT IN THE PROTEIN OF THE PRESENT INVENTION

1 mpasllqp rammktlssg nctlsvpakn syrmvvlgas rvvkssivsr flngrfedqy
61 tptiedfhrk vynirgdmyq 1dildtsgnh pfpmamrrlsi ltgdvfifvf sldnresfde
121 vkrqlqkqile vksclknktk eaaelpmvic gnkndhgclc rqvptteael lvsgdencay
181 fevsakkntn vdemfyvlfs maklphemsp alhrkisvqy gdafhprpfc mrrvkemday
241 gmvspfarrp svnsdlkyik akvlregqar erdkctiq

>CRA|18000005238449 /altid=gi|5059122 /def=gb|AAD38928.1|AF134409_1
(AF134409) Rhes protein [Rattus norvegicus] /org=Rattus
norvegicus /taxon=10116 /dataset=nraa /length=266
Length = 266

Score = 512 bits (1304), Expect = e-144
Identities = 252/266 (94%), Positives = 256/266 (95%)

Query: 1 MMKTLSSGNCTLSPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY 60
MMKTLSSGNCTL+VPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY
Sbjct: 1 MMKTLSSGNCTLNVPAKNSYRMVVLGASRVGKSSIVSRFLNGRFEDQYPTIEDFHRKVY 60
Query: 61 NIRGDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLDNRESFDEVKRLQKQILEVK 120
NI GDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLD+RESFDEVKRLQKQILEVK
Sbjct: 61 NIHGDMYQLDILDTSGNHFPAMRRLSILTGDVFILVFSLDSRESFDEVKRLQKQILEVK 120
Query: 121 SCLKNKTKEAAELPMVICGNKNDHGELCRQVPTTEAELLVSGDENSAYFEVSAKKNTNVD 180
SCLKNKTKEAAELPMVICGNKNDH ELCRQVP EAELLVSGDEN AYFEVSAKKNTNV+
Sbjct: 121 SCLKNKTKEAAELPMVICGNKNDHSELCRQVPAMEAELLVSGDENCAYFEVSAKKNTNVN 180

Query: 181 EMFYVLFSMAKLPHEMSPALHRKISVQYGDAFHPRFCMRRVKEMDAYGMVSPFARRPSV 240
EMFYVLFSMAKLPHEMSPALH KISVQYGDAFHPRFCMRR K AYGMVSPFARRPSV
Sbjct: 181 EMFYVLFSMAKLPHEMSPALHHKISVQYGDAFHPRFCMRRKVAGAYGMVSPFARRPSV 240

Query: 241 NSDLKYIKAKVLREGQARERDKCTIQ 266
NSDLKYIKAKVLREGQARERDKC+IQ
Sbjct: 241 NSDLKYIKAKVLREGQARERDKCSIQ 266

Hmmer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	126.2	2.8e-36	1
CE00060	rab_ras_like	20.7	0.00013	2

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
CE00060	1/2	22	52 ..	26	56 ..	16.9	0.0014
CE00060	2/2	162	183 ..	159	181 ..	3.3	7.5
PF00071	1/1	21	186 ..	1	169 [.	126.2	2.8e-36

1 CTCTCTGACT CTTTGCCTCC TCTCTGACTC CCTGCCTCCT CTCTCTGTCT
51 CCCTGCCTCC TCTGCTGAC TCCCTGCCTC CCCCTCTGT CTCACTGCCT
101 CCTCTCTCTG ACTCTCTGCC TCCTCTCTCT GACTCCCTGC CTCCCTCTCTC
151 TGATTCCCTG CCTCTTGAC CCTCTGCCTC CTCTCTTGA CTCCCTGCCT
201 CCTCTCTCCG ATTCTCTGCG TCTTGTACTC CCTGCCTCCT CTCTCTGACT
251 CCCTGAAGCT CATTCACTGCA TTGCTATCAA CTGCTGTGA CCAAGCTCTA
301 GGCTGGAGGC TGGGCAGGGC AATGATGGAG ACAAAACTCTG TCCCTGGGAG
351 CTTCTGGCCC CTTTCCCATC CTGTTAGAC AGAAGTGACC GCCAGCAGAG
401 TCAAGCTGTC TGCAAGAGGA CTTGGGGAGG GGGCTGTCT GGGTAGGGC
451 TTCTTCCCCC CCATCTCTGC TGAAGGCCCA GGCTGGCTGA GACAGCCCCG
501 GCAGAGACTG AGAAGGGCTC CCTGCTGTGG TCTGGCAGCC CCCTCTCCAC
551 CCTCTCTCT CTCATTTCT GCCTCCCACA CGTATGCCCT GGGCACCTCA
601 TCAGGGCTGC CCTAGGGGAG GGCCCTCCTT GCACAGGCC CTGGGCCAGT
651 CAGGTGGTT AGGCTGAGGA GAGAAGGTCC CAGAGTGGGG CTTCAGGCAA
701 ACCCAAAGAC AGAGCCCTTT GCCATTGAT GAATGCACAG ACCCTTTATT
751 GAGCCCCCTGC TCTGTTCATG GCATGGCAGT TTGTTGGGAT AAATTCAAG
801 ACAGCTTTAG GTGGGAGCTG GTGGGGGAT GTGGGGGTCT TAGGCTTGA
851 CTACTACCCA GCCTCTTTG TTAACCAAGT AGCTAGTCAC GTAGCCTCT
901 GAGCTGGGG CAGACCACT GGGATCAAAC CTCTCCTCTG CTGGTTACTG
951 GCTGTGCAAC TGTAAAG TAATTAAAC TCTCTGTGCC TCAGTTCTCT
1001 CATCTGTAAA TTGGAGATA ACACCACTG CTTTCTGGGG TTATGAAGGG
1051 AGAAATAGGT TACATGTT GCAGCACTTA GAACACTCTG GCATATTCTA
1101 GCTGAAAT GATGCCAGC TATGATTATT TCTATACTTA GTGCGGGGCT
1151 TGGCACACTG CATGGCTCA AGTGGCAGCA GTTGTGCTCC TTGTTGGCTCC
1201 AGGCTGGGG TCCCGCTGT GCTGAGCTGG CTTATTGTGC ACGTCCCTT
1251 GTGATTCTT CATCGAAGTC ACATTAGTAG CTTAGAAGTG ACCGTAGTGG
1301 GAGCATTAC GCATGGAAA TTGCAATAG GGCTTTAAC AAAGTATT
1351 TTGAGAGCCG GTTTCCTGCA CAGAGGCTGG TAGTTGGCA GGGTGAGCAG
1401 ATCCAGATGT GTGCCAGGG CTCGACGCA GGCAATCTCT CCACCTCCAG
1451 TGGCCATCTC AGACCTTAGC TTCACTGATAG CCAGGAAGCG ATGGTGTGG
1501 AAAGCGCTT GGGTCAATGG GCGAGGGCACT CAAGGAAACC GACTTGGGGC
1551 ATCCTGGGGT GGGGACCGAG TTTGGGCACA TACAGCCCTT TGTGTGAATT
1601 TAAAAAACAGT GCCTTTCTCT CTACACAAAGA TGCCCTTTCG TCTGGGATAC
1651 AGCCCCCACC TCTGGATGC AGCCCCCACT TGCCCACCCA GCCATGCGCC
1701 TTGTGCAAGTA TCCAACCTGC ACAACCTGTG GCAGCCTGTG GAAGACCGAG
1751 GGGATTGATA TTTCAGCAGG CCTGTGCCA TTTGCAGTTG AGGGGCTGGA
1801 AAGCTCTCCT CTGGAGAGGG GAGGGATTCC TGCAAGGGTG AGGAGATCAG
1851 AGAGGCCTTC AGAGAGCAGG TGGCACTTGA GCCAGACCC GAAACATAAG
1901 GGGAGAGGG TGTTCCTGAG AGGGGTGGCA TGAGCAAAGG AGTGGAGGCT
1951 GATCTCAGCA GAGCTAAAC TGACCGAGGGT GACTGGGGTC AGGGGTTCTG
2001 GGGCGGGGAT TCTGGTGGGC GCTAAGGTAG GAAAGGAGGG AGGGCTGGC
2051 TGTGAAGAGC CTTTGGGGT AGCCTGGTGG AGCCTGCGG TTTGCTATA
2101 CAAGAGCTT GATCCATGTC GGCTCTTTTCT ATGAGGTCAA GAGGCTCCCA
2151 TAGAAAGCTC TGAGTTTGCC CCAGAACCAT ACCCTTGGA GATGGGAGGG
2201 AAGCTTGAGC CAGCCATGGG TCGTCTCCCA TTCCACATCC TCTACTCCGG
2251 GCCCTGGGGT CTCTGGAGG CAAGTAAACA CCTAGGGCCT GGGAGGAAA
2301 AATATCCGGG CAGGTCACTGG AGCGAGGGGA GCCCCGCCAGA TGCAAGGAC
2351 AGGTCTAAAG GTGGGCTCTC CTGAGGTGGC TGCAGGAGCA ACCCCAGGCA
2401 TTGGGCTTGG AGCATGGCT GTGGACATAG CCTTCCCTTC TTCCCGAGGAG
2451 GGCTGAATGG CCACAGAAC ACCCCCTGGC CCAGGCTTAA GAAATGCATG
2501 CTAGTGCCTT CCCCCATGCT TATCCTAGAA TCACAGGCTC CGGGAAAGCC
2551 AGATGGATGA ACCAGGGAAA GAACGGATTG TCACCATAGA TACCATTTT
2601 GAGATTTCAC CATGTGCTGA GCCCTTGCA ACAACTCTAT GAAATGGGCT
2651 CATTGGCAG ATGAGAAAAG TGACTTCTAG AGAGGTTAAG CTACTAGCCC
2701 AAGATCAGTA GCTAGAGGCA AGGCAAGGAT TCAAACTCCA GGAGTCGGT
2751 GCTTGCATAA ATGAAAGGAT GAATGAACGG ATATTGAGTG AGTGAGTGG
2801 TGAAGGAAGG AGTAAAGGAG AGGGCATGAA TGAATGAGAG GGTAGAACTC
2851 CAAGACCCCT TAGAACCTCG TCTGATGTT CCATTTTACA GACAGAAAAC
2901 TGAGTCCTAG ACAGAGGCT AGAGGAGGCC AGAGGTTGGT GGGGCCAGGT
2951 CGGGGGGGCC CTGATGCTG CTTCTCTCGC TTTGTTGCA CCCCAGGCCA
3001 TGATGAAGAC TTTGTCCAGC GGGAACTGCA CGCTCAGTGT GCCCAGCAAA
3051 AACTCATACC GCATGGTGGT GCTGGGTGCC TCTGGGTGG GCAAGAGCTC
3101 CATCGTGTCT CGCTTCTCA ATGGCCGCTT TGAGGACAG TACACACCA

3151 CCATCGAGGA CTTCCACCGT AAGGTATAACA ACATCCGCGG CGACATGTAC
3201 CAGCTCGACA TCCCTGGATAC CTCTGGCAAC CACCCCTTCC CCGCCATGCG
3251 CAGGCTGTCC ATCCTCACAG GTGAGGCCCA CTGGTGCCTG GGCTGGGCG
3301 GCAGGGCCAG GGCATGGGT CGGAGTGTGC TGGGCACTTG GCAGTTGCA
3351 TAGACTTGCA TAGGCCATCGT CTGAGACAGG CGTCATCCCT GCACAATGAG
3401 GCTCAGAGAG GTTTGCAT GTGCTGGAAA TAGTGTGAA GTCGGGGGCC
3451 CCGATTCCAT TCTGTTAGAC TCCAGATCGA TTACTCATGG CTGTCGGGGC
3501 CGCCTTCCAG ATCAGGAGCT GATACCAGCA TGCCCCAGGG ATATTCCTTT
3551 CTAGGGAAACA GAATGATGCC CTGGCTGCTG CTTTCCTTCT CCGGAAGATG
3601 ACCCACCAGA GCTCCAGGGC CCAAGGTCA TCCACGGGGC TCAGGTCCTC
3651 CACACCCAG GCCTTGGCA CCTCTTAGAG AGGTAAAGGG AGGACCCAGG
3701 CAGTGTACAC CAAAGGGAAG GGGGCTTGGT CATGGTCATA GTGATGGTGA
3751 TGGCACTAGC TGACACTTAT CAGAACGCTAT GGGCCTGGCC CTGTTCTTAG
3801 AGCTTGGCAT GTAGTTTTT TTGAAACAGA GTCTCGCTCT GTCAACCCAGG
3851 CTGGAGTGCA GTGGCGCAT CTCGGCTCAC TGCAACCTCT GCCTCCCGGG
3901 TTCAAGCGAT TCTCCTGCC CAGTCCCCCA AGTAGCTGGG ACTACAGGCA
3951 CGTGCCACCA TGCCCGGCTA ATTTTTTGTA TTTTTACTAG AGACGGGATT
4001 TCACCATGTT AGCCAGGATG GTCTCGATCT CCTGACCTCG TGATCTGCCT
4051 ACCTCAGCCT CCCAAAGTGC TGAGATTACA GGCGTCAGCC ACCGCGGCCA
4101 GCGAGCATGT AGTTATTTAA CCCTCACAGT AAATAGTTAT TCATTCCCTT
4151 TTACAGGTG GGGAAACTGA AGCCCCAGAGA GTTAACTGAA CTCACCTCAG
4201 TGTTAGCACA GCTCGTAAAG GCAGTCTGCT TTTGTCCTT CAGACAAAGC
4251 CATACCCACAG CCTCTCAGCC CTGCTGGGAA GGTTGAGGAG GGACAGGGAG
4301 GTTGGGGGGAGA AGAAGGGGTG AGTGGAGCTG AGGGGCTGTG CCCTTGTTTA
4351 CACTGCAATTA GCATGGTGAAGA TAAGGAGGACA AGCCCCGGGAC CCAGCACCTG
4401 GGTGTGAGGCC CTGGTTCGGC TGCTTCCCTGG CTTTGTACCT CGAGGCAAGG
4451 GATTTTATCT CCTTGTGCT CAGCACTCTC ATCTGTAAAGA CTGACCCACA
4501 TCAACACTCA TCCTAAAGGG ACTGTGAGAC TTAAATGAAT GAATATATGT
4551 AAGGCGCTTG GTGAGCAGAT AGTAAATGCA CAATAAATCC CCAAGTCTTC
4601 TGTTGAGTCA GCATTTGCAA GTGGGCCTGC TACGGGTTAC ACGATCATT
4651 CCAAAGTCAAC GCCCCCTGAAG TTGCTGAGCA GGGATAAGGG AAGGAGTGAG
4701 CAGGCAACTC TCTAGGCATC ATTCAAGATAA CCCCCCAACT GAGGTACTTC
4751 TATACAGAGA AACCCATGCC ACTCCCAGCC CTGCTGCCGC CTTGAGGCCA
4801 AGACTGAGGC TGCGGGGTGG CCCCTCCTTG AGTGTCTTCT CTTCCAGGCT
4851 GGCTTTCCG AGCATCTGAC CCAGACAGCA GTCAAGTTCT CCGCTCCACC
4901 CCGAGTTTTG GAGAAGGGGC ATGTAGATAG GAGAGCCCTG GGTGACACTG
4951 TGTTGCAATC CTTGCTGGGC CTCTTGGCTT AAATGTGTGA CCAGAGGCCAC
5001 ATGCATCTG TCTGAGTCTC AGCCTCCCCA GCCACACAGT GGGCTTAACC
5051 TCATACCCCG CAGGGAGGCT GTGAGGACTG CAAGAAGGCT TGTGGGGGAA
5101 GCTTCCAGCA CGTGACGGGT ATTGCAATTGG TGTCAGCTCC CCCAGCTTGT
5151 GGGAGGGGAC TGGGTACCCG CTGCAATGAA TAAGGCTAAT GACAGAGGGAA
5201 AGGAGAGGGG AGATGTAGAG AGGAAGCACA TGCATATTTC CAGCATTAA
5251 TTTCAGTGAC ACAAGTAATA CCCCCAACACA CCCTCTGCA AACGCTACAG
5301 ATAAAGCTAA TGCCCCTTG ACCCATGTCC CCAATCCCA GCTCTGGCC
5351 CTGCCCGGA GGTGGCCACC CTGGCAGTCT GGCATGGAGC CTTCCGGGCC
5401 TCCGTGACTA CACCGGCATT CGTATTGTA TCCCCACAAT GGAGAGTATT
5451 TTGTCCTGTC TCTTTTTAT GGCAGCATATC ATTCTGAGCA CAGCTGCTG
5501 ATGCTTGTGTT TTTTTCTACA CACCAACCCCG TGCCTCATT TCCAACACTGG
5551 TGAACACTCA TTTTTCTAAC CTCATTTCC TGCTGCTCAG GAAATTCTGAA
5601 AAGCCATTAA TTCCACTGCC AGCTCTTCTT CCAGCTGCCA GACGGGCCGA
5651 TCTCTGATGC TTGGCATCGC AGTCTCGCAT TTGAATATGT CAAGGCCACG
5701 AGTCCTCAGG GGGCCCGGAT TAGCCTAACG GGATGGGGTT TGGCAGCCCA
5751 AGCAGGAGAAG GTGCCAACG TGACGCTGCC TGCGAAGTGC CTTTCAGAAG
5801 AGCCCCAACT GCAGTCTCCCT CCATCACCTC CCATCCATTG AGGCTTCTT
5851 GGTAAACACT GACTGTGTGC CAGGCCCTGG GGAGACCAAGG ACGAGTGGGT
5901 GATGGAACCC TTCTCTGTGC CCGAGCTGTT TGGAGCACAC CTTTGATCTG
5951 GACACCCATTG TGAATGTGCC ATGTGCCATT AAATGGGGGT AAATGATGTG
6001 CTCTGGGGGT GCAGAGGAAG GTGCCAGCCA TTCTGCCAGA AGCTGGAACCT
6051 GGTGCTTCT CTTCTCAAGA ATTGGGCCA ATTGCTGATT CCTCTGGGCC
6101 TCAGTTCTCT CATCTGTGAG ACAGGGATCT TGTACACCA CAAGGCTATC
6151 AAGAGTTTGA GCAAAGTGG TTGGACGCGAG TGGCTCATGC CTGTAATCCC
6201 AGCTCTTGG GAGGCCGAGG TGGCAGATC TCTTGAGGTC AGGAGTCAAG
6251 GACCAGCCTG GCTAACACAG TGAAACACCG TCTCTACTAA AAAATACAA

0
1
2
3
4
5
6
7
8
9
-
=

6301 AAATTAGCCA GGTGTGGTGA TGGGCACCTG TAATCCAGT TACTCGGGAG
6351 GCTGAGGCAG GAGAATCTCT TGAACCCAGG AGGTGGAGGT TGCAGTGAGC
6401 TGAGATCTTG CCATTGCATT CCAGGCTGGG CAACAAGAGT GAAACTCTGT
6451 CTCAGAAAAT AAATAAATAA ATAAAATAAA AATAGCTAGG CATGGTGACA
6501 GGCGCCTGTA ATCCCAGCTG CTCCAGAGGC TGAGGCAGGA GAATCGCTTA
6551 AACCCAGGG AGTGGAAAGTGC CAGTGAGCCA AGATCACACC ACTGCACTCC
6601 AGCCTGGGCC ACAGAGAAAG ACTCCATCTC AAAAAAAA AAAAAAAA
6651 AAAAGTTTA AGCAAAAGTG AGGAAGGTGC TTATTAAAAG CTGGAAATCA
6701 GGATGGAGGT ACCAGTCCAG ACAGCCTCCC CACCACCCA CCGTCTCCAC
6751 AGCAGCCCT GTTTCAGATT CACAAGCCTG CCTTGAGTGA TGCAGTGAGT
6801 TATCCTGGAG GCAGTGTGGG CCTTGGAGGC CAGCACTCAC TTTTCATCC
6851 TATGATTAT TTGAGAAGCA GAGAGCACCT ACCGGGTGCC AGGAACGAGC
6901 TAGGTGAGAA CAGAATCAGG TAGAAATCTC AGCCTAGCCA CACGGAGCT
6951 GTGTGATCTT GGGCAGGCTG CATAACCTTT CTGAGCCTCA GTTTGCTCAC
7001 CTGTAATGCA AAGGTAACAA ATCTTGACA GAGGCATAGT GAGGAATCAA
7051 GAGAACAAACG GGCCTGGAGC ATACACCCAG TGCTTAGCCC CCAGTAGGCC
7101 CTCACTCTCA TCATTACTGA CACCTGAGGT CACTGAGCAT GTGCCACTGT
7151 CCATTCACTA TCTTGATCAA CTCCCCAAAT CATCCTGCAA GGTAAATATT
7201 CATCTTCATG AAACAGACAG AGAAACTGAG GTTACAGAGG TTTCTGTGATC
7251 TGCCCAAGTC TGCTGGCAGC TAAGCGGATG AGGCCAGATG CAAACTAGGC
7301 ATTGAGCAAG ACAGGCAGGA CCCCTGCTCT CATAGAAATG ATTTTTATTA
7351 TTATCTGAAC ACAGTCCACA CAAGTGCACCT ACCCCTCTCC AGCCCTGCAA
7401 AGAAATGTGA AGTGAGTTAA CTGTATTTGA ACCAAGTGGT CCACGTGTTA
7451 GCTATGCGAC TGTGAACAGG GGCTTCACCC CCCTCAGCCT CAGTTCTGT
7501 TCCTGGAAAA TAATCGCAGG GAGAATAATC GCAGCTACCC CGAAGAGTCG
7551 CTGTGTAGGT TAAAGCAGTT ATGCCGCATA ACTGCTTCAG GCCACCTGTG
7601 ACTCCCAGCT CTTAGGGCTG ATGTTCTGTG GCCAGAGGAG GGCAGGGGTT
7651 GCAGCTGGCC GGTGAACTCA CTACCTGGGC TCTCTCCCTG CAGGGGATGT
7701 CTTCATCCTG GTGTCAGCC TGGATAACCG GGAGTCCTT GATGAGGTCA
7751 AGCGCCTCA GAAGCAGATC CTGGAGGTCA AGTCCCTGCCT GAAGAACAAAG
7801 ACCAAGGGAGG CGGGGGAGCT GGCATGTC ATCTGTGGCA ACAAGAACGA
7851 CCACGGCGAG CTGTCGGGCC AGGTGCCAC CACCGAGGCC GAGCTGCTGG
7901 TGTCGGGCGA CGAGAACTGC GCCTACTCG AGGTGTCGGC CAAGAAGAAC
7951 ACCAACGTGG ACGAGATGTT CTACGTGTC TTCAGCATGG CCAAGCTGCC
8001 ACACGAGATG AGCCCCGCCA TGCTATGCAA GATCTCCGTG CAGTACGGTG
8051 ACGCCTTCCA CCCCAGGCC TTCTGCATGC GCCCGCTCAA GGAGATGGAC
8101 GCCTATGGCA TGGTCAGCC CTTCGCCCG CGCCCCAGCG TCAACAGTGA
8151 CCTCAAGTAC ATCAAGGCCA AGGTCCCTCG GGAAGGCCAG GCCCGTGAGA
8201 GGGACAATG CACCATCCAG TGAGCGAGGG ATGCTGGGAG GGGGCTTGGC
8251 CAGTGCCTTC AGGGAGGTGG CCCCAGATGC CCACTGTGCG CATCTCCCCA
8301 CCGAGGCCCG GGCAGCAGTC TTGTTCACAG ACCTTAGGCA CCAGACTGG
8351 GGCCCCCGGG CGCTGGCCTC CGCACATTG TCTGCCCTCT CACAGCTTTC
8401 CTGAGTCCGC TTGTCCACAG CTCCCTGGTG GTTTCATCTC CTCTGTGGG
8451 GGACACATCT CTGCAGCCTC AAGAGTTAGG CAGAGACTCA AGTTACACCT
8501 TCCCTCCCTG GGGTGGAAAG AAATGTTGAT GCCAGAGGG TGAGGATTGC
8551 TGCCTCATAT GGAGCCTCTC GGGACAAGCC TCAGGATGAA AAGGACACAG
8601 AAGGCCAGAT GAGAAAGGTC TCCCTCTCTC TGGCATAACA CCCAGCTTGG
8651 TTTGGGTGGC AGCTGGGAGA ACTTCTCTCC CAGCCCTGCA ACTCTTACGC
8701 TCTGGTTCAAG CTGCCTCTGC ACCCCCTCCC ACCCCCCAGCA CACACACAAG
8751 TTGGCCCCCA GCTGCGCCTG ACATTGAGCC AGTGGACTCT GTGTCTGAAG
8801 GGGCGTGGC CACACCTCTC AGACCACGCC CACCACTTAG ACCACGCCA
8851 CCTCCTGACC GCGTTCTCA GCCTCCTCTC CTAGGTCCT CCGCCCGACA
8901 GTTGTGTTT GTTGTGGTT CAGCTGTTT CGTGTGATGT ATAGTAGTAG
8951 AAATGGAAAT CATTGACTG TAAAAGCTA GTGACTCCCT CCTTGGCCAG
9001 GCCCTCACCC AGTTCAGATC CACGGCTCC ACCCGGGACG CCTTCCCTCCT
9051 CTGCTCCCAA ACAGGGTTTC CGTGGCCTGT TTGCAGCTAG ACATTGACCT
9101 CCGCCATTGA GCTCCACGGT TTACAGACAA TTGCACAAGC GTGGGGTGGG
9151 CAGGCCAGGA CTGCTTTTT TTAATGCTCC CATTTCACAG AGGATACAC
9201 CGAGACTCGG AGGGGACACG ATGAGCACCA GGCCCCACCT TTGTCCTCA
9251 GCAAATTCAAG GGTACAGCTC CACCTAGAAC CAGGCTGCC TCTACTGTGC
9301 TCGTTCCCTCA AGCATTATT AAGCACCTAC TGGGTGCTGG GTTCAGTGTG
9351 TCCTAGGAAA CCAAGAGGGT CCCCAGTCTT GGCCTCTGCC CGCCCCCTGCT
9401 GCCCCACCAAC CTTCTGCACA CACAGCGGTG GGGAGGCAGG GAGGAGCAGC

9451 TGGGACCCAG AACTGAGCCT GGGAGGGATC CGACAGAAAAA GCTCAGGGCG
 9501 GGTCTTCCTCC TTGTGCCCGG GATTGGGCTA TGCTGGGTAC CACCATGTAC
 9551 TCAGGCATGG TGGGTTTGA ACCCATAAAC CAAAGGCCCT TGTCTACAGC
 9601 TCTTAACAAG TATATTTGT ATTTTAATCT CTCTAAACAT ATTGAAGTTT
 9651 TAGGGCCCTA AGGAACCTTA GTGATCTTCT ATTGGGTCTT TCTGAGGTT
 9701 AGAGAGGGTA AGTAACCTCC TCCAGGTAC ACAGCAAGTC TGTGGGTGGC
 9751 AGAAGCAAGC TAGCGCTGGG CATTCACTAC ATACCAAGAT GTGCTCCCTC
 9801 TCTTGATGCT TGGCCCCCTGG GCCCTTCAGG GCTTGGGAC ATCTTGTCT
 9851 CAACCCCTCTC CCTAGATCAG TCTGTGAGGG TCCCTGTAGA TATTGTGTAC
 9901 ACCATGCCCA TGTATATACA AGTACACACA GATGTACACA CAGATGTACA
 9951 CATGCTCCAG CCCCAGCTCT GCATACCTGC ACCTGCACCC CAGCCTTGGC
 10001 CCCTGCCTGC GTCTGTGCTC AAAGCAGCAG CTCCAACCCT GCCTCTGTCC
 10051 CCTTCCCCAC CCACTGCCTG AGCCTTCTGA GCAGACCAGG TACCTTGGCT
 10101 GCACCGGTGT GTGGCCCGCT CTCACCCAGG CACAGCCCCG CCACCATGGA
 10151 TCTCCGTGTA CACTATCAAT AAAAGTGGGT TTGTTACAAA GCCGTGTCT
 10201 TGCCCAGTGT TATTTTTGT ATTTCCAAGA GGAGGTGTGC CCCTTTCCAG
 10251 ACCAAAGCTG GCCTTCCCT CCCAAATGC ACCTGCCGTG TACCTTGGCC
 10301 CTGAGGGTCA GCACTGAGTC CACCTCAAG TGTAAGTGTG GGGAGAGGGG
 10351 GATAAGTCCC CCAGATGGAA GGTGATGCC TCCTTCAGCC TGGCCCTCCT
 10401 GGGTCCTCCG GGTGTGTGTA CCGAGGTGTC TGTGTCCACA AAGAAGGGC
 10451 CCCCGTGGAC CATTAGCTCC AGGAGGATCT CCGTGTCTGA GTTCTTTGTG
 10501 ATTCCTGTAC AGCAGCAATT TCACCCCGAG GGGACAGTTG GCAATCTCTG
 10551 GAAACCTTT CCAAGCCTGG GGCTGGGGCT GCTACTCTCA TCTGGTGGGT
 10601 GGAGGCCAGG GACACCATTG AGTATCCTCC AACGCACAGG ATGCCCTCCT
 10651 ACCCCCCACCC CACTGAGAAT TATCTGGCCT CAAATGCCA GCGTGGGCAG
 10701 CCTTACTTAG ACTCACCCCA GGGGCTGGGA CACGCCCTCA CCTGCCTGTG
 10751 ATGGATTGTG TGGACCACAT TCTGGACGGA ACCCACAGCA TAAGCACTCC
 10801 TGTGAAGTGA GACAGGATGT GGGTGAGGAT GGAAAGTGGA GGCTGAGGG
 10851 GAAGGTCCTGG GCCCTGACCA ACACGGAATG TGCCCCCTGG GACTGAGAGG
 10901 CTTCCCTGGG CAGAGGGAAA GGAGGAAGTC AGTGAAGTAA AATACTCCCT
 10951 GTGTGTTTA CCCAGCGAGT CTCACGCCAT CCTATCACCC AGCCCCGAGG
 11001 GAAGCCCACT CATGTTCACCC CCATCTGAGC ATTTAGGCTC AGAGAGCTCA
 11051 ATATCTGTGTC CAAGATGGCA CAGCTGGTGA AGTGGCAGAT CAGAGATTCA
 11101 ACACCAGAGG CTGTCGATT TCCGTCTGGC TGAAGAAAGA TTTTGCATCA
 11151 GGGAGGTGGA AACCATCTGT GCTTTTGATC AGCAAATGCC ACCAGCAGGA
 11201 TCAGGGAGCC AGGCCATAAA G

FEATURES:

Start: 3000
 Exon: 3000-3270
 Intron: 3271-7693
 Exon: 7694-8220
 Stop: 8221

CHROMOSOME MAP POSITION:

Chromosome 22

ALLELIC VARIANTS (SNPs):

Position	Major	Minor	Domain
3951	C	T	Intron
4127	C	T	Intron
4157	G	A	Intron
4513	C	T	Intron
6894	C	A	Intron
8409	G	C	Beyond ORF(3')
8437	T	G	Beyond ORF(3')
8579	T	C	Beyond ORF(3')
10292	A	G	Beyond ORF(3')
10792	A	G	Beyond ORF(3')

Context:

DNA Position	Sequence
3951	CACACCCCAGGCCTTGCCACCTCCTAGAGAGGTAAGGGCAGGACCCAGGCAGTGATCAC CAAAGGGAAGGGGGCTTGGTCATGGTCATAGTGATGGTATGGCACTAGCTGACACTTAT CAGAAGCTATGGGCCTGGCCCTGTTCTAGAGCTTGGCATGTAGTTTTTGAAACAGA GTCTCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCATCTCGGCTCACTGAAACCTCT GCCTCCGGGTTCAAGCATTCTCCTGCCTCAGTCCCCAAGTAGCTGGACTACAGGCA [C, T] GTGCCACCATGCCGGCTAATTTTGATTTTACTAGAGACGGATTTCACCATGTTA GCCAGGATGGTCTGATCTCTGACCTCGTATCGCTACCTCAGCCTCCAAAGTGC GAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTAACCCCTACAGTA AATAGTTATTCACTCCCTTTTACAGGTGGGAAACTGAAGCCCAGAGAGGTTAAGTAAC TCACTCCAGTGGTAGCACAGCTCGTAAAGCAGTCTGTTTGTGCTTCAGACAAAGCC
4127	CAGAGTCTCGCTCTGTCACCCAGGCTGGAGTGCAGTGGCGCATCTGGCTCACTGCAAC CTCTGCCTCCGGGTTCAAGCGATTCTCTGCTCAGTCCCCAAGTAGCTGGACTACAA GGCACGTGCCACCATGCCGGCTAATTTTGATTTTACTAGAGACGGATTTCACCA TGTTAGCCAGGATGGTCTGATCTCTGACCTCGTATCGCTACCTCAGCCTCCAAA GTGCTGAGATTACAGGCGTCAGCCACCGCGGCCAGCCAGCATGTAGTTATTAACCCCTCA [C, T] AGTAATAGTTATTCACTCCCTTTTACAGGTGGGAAACTGAAGCCCAGAGAGGTTAAG TAACTCACTCCAGTGGTAGCACAGCTCGTAAAGGCAGTCTGCTTTGTGCTTCAGACAA AGCCATACCAACAGCCTCTCAGCCCTGCTGGGAAGGGTGGAGGAGGGACAGGGAGGTTGGGG GGAAGAAGGGGTGAGTGGAGCTGAGGGCTGTGCCCTGTTACACTGCAATTAGCATGGT AGCTAAGAGGACAAGCCGGGACCCAGCACCTGGGTGTGAGCCCTGGTCCGCTGCTTCC
4157	TGCAGTGGCGCGATCTGGCTCACTGCAACCTCTGCCTCCGGGTTCAAGCGATTCTCCT GCCTCAGTCCCCAAGTAGCTGGACTACAGGCACGTGCCACCATGCCGGCTAATTTT TGTATTTTACTAGAGACGGGATTTCACCATGTTAGCCAGGATGGTCTGATCTCTGAC CTCGTGTCTGCCACTCAGCCTCCAAAGTGCTGAGATTACAGGCGTCAGCCACCGCG GCCAGCCAGCATGTAGTTATTAACCCCTCACAGTAAATAGTTATTCATTCCCTTTTACA [G, A] GTGGGGAAACTGAAGCCCAGAGAGGTTAAGTAACACTCCAGTGGTAGCACAGCTGTA AAGGCAGTCTGCTTTGTGCTTCAGACAAAGCCATACCACAGCCTCTCAGCCCTGCTGG GAAGGGTGGAGGAGGGACAGGGAGGTTGGGGGAAGAAGGGGTGAGTGGAGCTGAGGGGCT GTGCCCTGTTACACTGCAATTAGCATGGTAGCTAAGAGGACAAGCCGGGACCCAGCAC CTGGGTGTGAGCCCTGGTCCGCTGCTTGTACCTCGAGGCAAGGGATTTTA
4513	TCGTAAAGGCAGTCTGTTTGCTTCAGACAAAGCCATACCACAGCCTCTCAGCCCT GCTGGGAAGGGTGAGGAGGGACAGGGAGGTTGGGGGAAGAAGGGGTGAGTGGAGCTGAG GGCTGTGCCCTGTTACACTGCAATTAGCATGGTAGCTAAGAGGACAAGCCGGGACCC AGCACCTGGGTGTGAGCCCTGGTCCGCTGCTTCTGGCTTGTACCTCGAGGCAAGGGAA TTTATCTCCTGTTGTCAGCAGTCTCATCTGTAAGACTGCACCAACATCAACACTCATC [C, T] TAAAGGGACTGTGAGACTTAAATGAATGAATATATGTAAGGCCTTGGTGAGCAGATAGT AAATGCACAATAATCCCCAAGTCTTCTGTTGAGTCAGCATTTGCAAGTGGGCTGCTAC GGGTTACACGATCATTCCCCAAGTCACGCCCTGAAGTTGCTGAGCAGGGATAAGGGGAAG GAGTGAGCAGGCAACTCTCTAGGCATCATTAGATAACCCCCCAACTGAGGTACTTCTAT ACAGAGAAACCCATGCCACTCCCAGCCCTGTCGCCGCTTGAGGCAAGACTGAGGTGTC
6894	GCACCTCCAGCCTGGCCACAGAGAAAGACTCCATCTAAAAAAAAAAAAAAAAAAAA AAGTTAAGCAAAAGTGAGGAAGGTGCTTATTAAGCTGGAAATCAGGATGGAGGTAC AGTCAGACAGCCTCCCCACCAACCCACCGTCTCCACAGCAGCCCTGTTCAGATTCA AAGCTGCCTGAGTGATGCAGTGAGTTATCTGGAGGCAGTGTTGGCCTTGGAGGCCAG CACTCACTTTCATCTATGATTATTTGAGAAGCAGAGAGCACCTACCGGGTGCAGG [C, A] ACGAGCTAGGTGAGAACAGAACAGTCAAGGTAGAAATCTCAGCCTAGGCCACACGGAAAGCTGTT GATCTTGGCAGGCTGCATACCCCTTCTGAGCCTCAGTTGCTCACCTGTAATGCAAAGG TAACAAAATCTTGACAGAGGCATAGTGAGGAATCAAGAGAACAGGGAAAGGAA ACCCAGTGCCTAGCCCCAGTAGGCCCTCACTCTCATCATTACTGACACTGAGGTACT GAGCATGTGCCACTGTCCATTCAATTATCTGCAAAATCATCTGCAAGGTA

8409	CATGGTCTGCCCTCGCCGCCGCCCCAGCGTCAACAGTGACCTCAAGTACATCAAGGC CAAGGTCTCGGGAAAGGCAGGGCCCGTGAGAGGGACAAGTGCACCATCCAGTGAGCGAG GGATGCTGGGGCGGGCTTGGCCAGTGCCTCAGGGAGGTGGCCCCAGATGCCACTGTG CGCATCTCCCACCGAGGCCCGCAGCAGTCTGTTCACAGACCTTAGGCACCAGACTG GAGGCCCGGGCGCTGGCTCCGCACATTGCTCGCTTCACAGCTTCGAGTCC [G, C] CTTGTCCACAGCTCCCTGGTGGTTCATCTCCTCTGTGGGAGGACACATCTCTGCAGCCT CAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTCTGGGGTTGGAAGAAATGTTGA TGCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCTGGGACAAGCCTCAGGATGA AAAGGACACAGAAGGCCAGATGAGAAAGGTCTCTCTGGCATAACACCCAGCTTG GTTTGGGTGGCAGCTGGGAGAACTTCTCTCCAGCCCTGCAACTCTTACGCTCTGGTTCA
8437	AGCGTCAACAGTGACCTCAAGTACATCAAGGCCAAGGTCTCGGGAAAGGCCAGGCC GAGAGGGACAAGTGCACCATCCAGTGAGCGAGGGATGCTGGGGCGGGCTTGGCCAGTGC CTTCAGGGAGGTGGCCCCAGATGCCACTGTGCGCATCTCCCCACCGAGGCCCGGCAGC AGTCTTGTTCACAGACCTTAGGCACCAAGACTGGAGGCCCGGGCGCTGGCTCCGCACA TTCGTCTGCCTCTCACAGCTTCCGTAGTCCGCTTGTCCACAGCTCCTGGTGGTTCA [T, G] CTCCTCTGTGGGAGGACACATCTCGAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACA CCTTCCTCTCTGGGGTTGGAAGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTC TATGGAGCCTCTGGGACAAGCCTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAG GTCTCCTCTCTCTGGCATAACACCCAGCTTGGTTGGGAGCAGTGGAGAAACTTCTC TCCCAGCCCTGCAACTCTTACGCTCTGGTTCAAGCTGCCCTGCAACCCCTCCCACCCCC [T, C] GCCCACTGTGCGCATCTCCCCACCGAGGCCCGGCAGCAGTCTTGTTCACAGACCTTAGG CACCAAGACTGGAGGCCCGGGCGCTGGCTCCGCACATTGCTGCGCTTCACAGCTT TCCTGAGTCCGCTTGTCCACAGCTCTTGTGGTTCATCTCCTCTGTGGAGGACACAT CTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTGGGGTTGGA AGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCTGGGACAAG [T, C] CTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCTCTCTGTGGCATAAC ACCCAGCTGGTTGGGGCAGCTGGGAGAACTCTCTCCAGCCCTGCAACTCTTACG CTCTGGTTAGCTGCCTCTGCACCCCTCCCACCCCCAGCACACACACAAGTTGGCCCC AGCTGCGCCTGACATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCC TAGACCACGCCACCACTAGACCACGCCACCTCTGACCGCGTCTCTCAGCCTCTCT
8579	GCCCAGACTGTGCGCATCTCCCCACCGAGGCCCGGCAGCAGTCTTGTTCACAGACCTTAGG CACCAAGACTGGAGGCCCGGGCGCTGGCTCCGCACATTGCTGCGCTTCACAGCTT TCCTGAGTCCGCTTGTCCACAGCTCTTGTGGTTCATCTCCTCTGTGGAGGACACAT CTCTGCAGCCTCAAGAGTTAGGCAGAGACTCAAGTTACACCTTCCTCTGGGGTTGGA AGAAATGTTGATGCCAGAGGGGTGAGGATTGCTGCGTCATATGGAGCCTCTGGGACAAG [T, C] CTCAGGATGAAAAGGACACAGAAGGCCAGATGAGAAAGGTCTCTCTCTGTGGCATAAC ACCCAGCTGGTTGGGGCAGCTGGGAGAACTCTCTCCAGCCCTGCAACTCTTACG CTCTGGTTAGCTGCCTCTGCACCCCTCCCACCCCCAGCACACACACAAGTTGGCCCC AGCTGCGCCTGACATTGAGCCAGTGGACTCTGTGTCTGAAGGGGGCGTGGCCACACCTCC TAGACCACGCCACCACTAGACCACGCCACCTCTGACCGCGTCTCTCAGCCTCTCT
10292	ACCCCTGGCCCTGCCCTGCGTCTGTGCTCAAAGCAGCAGCTCCAACCCCTGCCCTGTCCC CTTCCCCACCCACTGCCCTGAGCCTCTGAGCAGACCCAGGTACCTGGCTGCACCGGTGTG TGGCCCGCTCTCACCCAGGCACAGGCCACCATGGATCTCGTGTACACTATCAATA AAAGTGGTTGTTACAAGCCGTCTTGCCTGCCATGTGTATTTTGATTTCCAAGAG GAGGTGTGCCCTTCCAGACCAAAGCTGCCCTTCCCAAATGCACCTGCCGTGT [A, G] CCCTGGCCCTGAGGGTCACTGAGTCCACCTCAAGTGTAAAGTGTGGGAGAGGGG TAAGTCCCCAGATGGAAGGTGATGCCCTCTCAGCCTGCCCTCTGGGTCTCCGG TGTGTGTACCGAGGTGTCGTGTCACAAAGAAGGGGCCCGTGGACCATTAGCTCCAG GAGGATCTCGTGTCTGAGTTCTTGTGATTCTGTACAGCAGCAATTCAACCGCAGGG GACAGTTGGCAATCTCGAAACCTTCCAAGCTGGGCTGGGCTGCTACTCTCATC
10792	TTCTTTGTGATTCTGTACAGCAGCAATTCAACCCGAGGGGACAGTGGCAATCTCTGG AAACCTTTCCAAGCCTGGGCTGGGCTGCTACTCTCATCTGGTGGGTGGAGGCCAGGG ACACCAATTCAAGTCCAAACGCACAGGATGCCCTCCACCCCCACCCCACAGCAT ATCTGGCCTCAAATGCCAAGCGTGGGCAGCCTTACTTAGACTCACCCAGGGCTGGGAC ACGCCACCTCTGCGTGTGATGGATTGTTGGACACATTCTGGACGGAACCCACAGCAT [A, G] AGCACTCCTGTGAAGTGTGAGACAGGATGTGGGTGAGGATGGAAGTGGAGGCTGAGGGAGA AGGTCTGGCCCTGACCAACACGGAATGTGCCCTCTGGACTGAGAGGCTCCCTGGGCA GAGGGAAAGGAGGAAGTCAGTGAGGTAAGGACTCCCTGTGTGTTTACCCAGCGAGTCT CACGCCATCTTACACCCAGCCCCAGGGAAAGCCACTCATGTTCACCCCATCTGAGCAT TTAGGCTCAGAGAGCTCAATATCTGTCCAAGATGGCACAGCTGGTGAAGTGGCAGATCA